REMARKS

Posture of the case

The present application was filed on October 31, 2003. Claims 1-13 were originally presented. An Office action of August 15, 2007 presented a restriction requirement. Applicant traversed the restriction requirement in a response of September 17, 2007.   In the same response, Applicant also amended all claims in order to more certainly ensure that the claims were directed to a single invention, to provide sufficient antecedent basis for all terms, and to properly format the claims.

A first Office action of December 12, 2007, rejected claims 1-13 under 35 USC 112, second paragraph, claims 1-13 under 35 USC 101, and claims 1-13 under 35 USC 103(a) as being unpatentable over Computer and Chemistry, 1999, Vol. 23, p.365-385 ("Taylor"), in view of US 5,832,272 ("Kalantery").

In Reply A of March 11, 2008. Applicant canceled claims 12-13, amended claims, 1, 2, and 3-6 to overcome the 112, second paragraph rejection, amended claim 1 to overcome the 101 rejection, and amended claim 1 and it's dependent claims to overcome the 103 rejection.

A final Office action of July 9, 2008, objected to the specification, rejected claims 2-11 under 35 USC 101, and rejected claims 1-11 under 35 USC 103(a) as being unpatentable over Taylor in view of Kalantery, and in view of Proteins: Structure, Function, and Genetics, 1991, Vol. 11, p. 59-76 ("Huysmans").

In an amendment accompanying a Request for Continued Examination of October 8, 2008, Applicant amended the specification to overcome the objection, amended claims 2-11 and traversed to overcome the 101 rejection, canceled claim 1 and added claims 14-23 to overcome the 103(a) rejection.

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

Present Office action

1. Objection to specification

The present, nonfinal Office action of February 10, 2009, maintains the prior objection to the specification on grounds that the amendment filed 10/08/2008 is not in compliance with the MPEP and 37 CFR 1.121, due to providing a replacement paragraph having no markings to indicate changes. Applicant respectfully submits that the previously submitted replacement paragraph was submitted in a manner wherein the existing paragraph was deleted and a clean new paragraph was submitted in its place, which complied with the following requirements cited in the present Office action:

> Amendments to the specification, including . . . an amendment of a paragraph, must be made by submitting: (i) An instruction, which unambiguously identifies the location, to. . . replace a paragraph with one or more replacement paragraphs . . . ; . . (iii) The full text of any added paragraphs without any underlining; and (iv) The text of a paragraph to be deleted must not be presented with strike-through or placed within double brackets" (emphasis added).

Nevertheless, Applicant recognizes that reasonable interpretations of the requirements can differ and herein resubmits the replacement paragraph in the manner directed by the present Office action.

2. New amendment to specification

In reviewing the present application, applicant has noticed certain disagreement between the descriptions of FIG's 4-6 in the specification and the figures themselves. That is, upon review, it is unambiguously clear that the specification indicated incorrect figure numbers for certain portions of the description. Applicant accordingly herein submits amendments to the specification paragraphs as set out herein above to correct references to figure numbers.

3. Rejection under 35 USC 101

The present Office action rejects claims 3-6, 10, 14-23 under 35 USC 101 on grounds that the claimed subject matter does not either (1) tie a claimed process to a particular machine, manufacture, or composition of matter, or else (2) recite a claimed

process that transforms a particular article into a different state or thing, citing In re Bilski, 88 USPQ2d 1385 Fed. Cir. 2008 and In re Comiskey, Fed. Cir., No. 2006-1286.

The Office action indicates that in the present case the claims are directed to steps for manipulating data and are missing the use of a specific machine for carrying out these steps. The Office action helpfully points out that to qualify as a statutory process, the claims may recite use of a machine within the steps of the claimed subject matter, whereas mere preamble limitations that recite a process comprising machine implemented steps are not sufficient if the process otherwise recites only mental steps in the body of the claim.

Applicant herein amends claim 14 to more clearly indicate how a computer system is active in performing aspects of the claimed process steps. Applicant submits that this overcomes the rejection. No new matter is added, since the original application provides support for the amendments. See, for example, original application, as published, paragraphs 100-108.

4. Rejection under 35 USC 103(a)

The present Office action rejects all claims under 35 USC 103(a). Specifically, the Office action rejects claims 4, 5, 10, and 14-23 as being unpatentable over Taylor in view of Schwartz et al., Genome Research, 2000, Vol. 10, p. 577-586 ("Schwartz"), and rejects claims 3-6 as being unpatentable over Taylor in view of Schwartz and further in view of Huysmans. Applicant herein amends independent claim 14 to overcome the rejection, amends claims 4, 5, 6, 15, 17, 19, 21 and 22 to conform these dependent claims to claim 14, and cancels claim 23.

A. Discussion of the present application

The present application discloses a new match-set data structure and associated processes that enable an easily changeable view of a sequence representation, i.e., that enable a user to interact with a computer system so that the system generates and presents to the user a new sequence representation based on new or previously unchosen replets, where the new representation does not require regenerating position parameters of the match-set entries for the other replets used in the sequence representation.

It is known to represent sequences using an ordered set of match-set entries and a backbone, where each match-set entry represents a subsequence that starts at a location "k" of the sequence. Present application, paragraph 54. The sequence may be a human gene sequence or other sequence. Present application, paragraphs 23 and 38. A known pattern discovery algorithm "TEIRESIAS" generates conventional match-set entries (as shown in Table 2) for replets that match subsequences in a sequence. Present application, paragraph 74. "As an example, consider the first entry in Table 2. This entry provides the information concerning the pattern `cgcgcgcgcg,` that is the sequence in which occurs (0) and the offset (19) of the occurrence." Present application, paragraph 75.

The present patent application, however, discloses multiple ways to represent such a sequence, depending upon which replets are chosen for the representations. Present application, paragraphs 23-25; also present application, paragraph 49 (". . . a subsequence may be matched by more than one replet. The choice among these replets is made . . ."); present application, paragraph 55 ("Whenever a subsequence could be represented by one or more replets or one or more combination of replets, a choice is made among them and only one among these is used to represent the subsequence."). See also, paragraphs 109-135 (discussing different applications and ways of viewing a sequence that are facilitated by the disclosure of the present application).

The replets not chosen "also have an entry in their match-set entries against the sequences, which enables processing based on these replets. " Present application, paragraph 56. But these entries would be invalid since their matching subsequence is represented by the replet that has been chosen and then the subsequence has been removed from the sequence, so that conventional Match-set entries for these non-chosen replets would not enable rebuilding of the sequence. Id. Therefore, the present application teaches that the match-set entries for each replet includes the conventional "k" position parameter and a modification, i.e., a "δ" position parameter that indicates the number of positions before or after the position k that the replet starts matching the subsequence. Id. See also, present application, paragraph 43 (The match-set of the present application "describes the positional information of the replet

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

110 in a sequence 105. A Match-Set is a set of <seq_id, k, δ > ensembles. The variable "seq_id" indicates the sequence 105 where the replet 110 has a match, the sum of "k, δ" provides the starting position of the subsequence (that matches the match-set's replet) in sequence seq_id. "). This change in the match-set enables mapping that would otherwise be difficult to perform. Present application, paragraph 57.

According to the issue explained above, non-chosen replets correspond to a subsequence but are at least initially not used to represent and reconstruct the overall sequence because another replet is chosen. Another issue also gives rise to replets that are at least initially not used to represent a sequence. That is, replets may be added for consideration (i.e., added to the "ontology") after match-set entries have been generated to represent a sequence. The data structure taught in the present application also allows adding replets after generating match-set entries for a sequence representation. Present application, paragraph 12.

The present application provides an example of how the new data-structure for match-set entries described therein operates. Present application, paragraphs 71-73 ("This example demonstrates how new replets are accommodated, and describes an algorithm and methodology for reconstructing the sequences from the data structures. Let the set of optimal patterns chosen to represent the set of sequences be $\Phi\alpha$ . . . Let the entire set of patterns chosen for representing the sequences be $\Phi+$ . . . Let the new replet to be added after the Replet-sequence matrix for $\Phi+$ is constructed be {actata}. . . . Table 3 below presents the information obtained by transforming the results in Table 2 above, generated using the TEIRESIAS algorithm, such that the information is structured in accordance with the required Match-Set datastructure. As an example, consider the first entry in Table 2. This entry provides the information concerning the pattern 'cgcgcgcgcg', that is the sequence in which occurs (0) and the offset (19) of the occurrence. The entries of Table 2 are modified to have k,δ parameters, and the resulting set of Match-Set entries as shown in Table 3 below.").

The foregoing summarizes how the present application discloses a new match-set entry data structure for replet position that is required to enable an easily changeable view of a sequence representation, i.e., a new sequence representation based on new or previously unchosen replets where the new representation does not

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

require regenerating position parameters of the match-set entries for the other replets used in the sequence representation.

The present application also discloses other new features in the match-set entry data structure that facilitate the easily changeable view. In particular, the present application discloses that the match-set entry data structure includes an Is-base-replet array, a Pointer to Base-replet array, a sequence-formation-edges array, a pointer to next-pattern instance, and a pointer to previous-pattern instance, as summarized in Table 7. The sequence-formation-edges array is a vector, such that the entry at index "i" represents the ith instance of the pattern on the sequence sequence-id, the matching-offsets array contains the various offsets at which the replet has matched the sequence, and the is-base-replet array indicates whether the replet was used to represent the sequence at that offset provided in matching-offsets array. Present application, paragraphs 97-98.

Formation of these arrays is illustrated schematically and discussed in the context of an example in the present application. Figure 4 in the present application illustrates, and paragraph 79 describes, a base replet-sequence matrix for the $\Phi\alpha$ example. See also present application, paragraph 45. (base replet-sequence matrix is the "replet-sequence matrix constructed using only those replets 110 that are used to represent a sub-sequence in a sequence 105."). FIG. 5 presents a replete-sequence-matrix 500 "that is modified to accommodate the overlapping pattern {aa..a . . . a} and the schematic representation of the resulting replet-sequence-matrix," where "[t]he base-replet-connector allows the resolving of the base pattern that was chosen against the non-base pattern (In this case, the pattern is {aaataa..aaa})." Present application, paragraphs 81 and 82, as amended herein. FIG. 6 presents a replete-sequence-matrix 600 "that is modified to include a new replet {actata}, and for which base-replet connectors are added from actata's replet instances to the corresponding tactata.....ttac's replet instances. Present application, paragraph 84, as amended herein.

The present application also discloses variation information that has to be stored if patterns $\Phi\alpha$ are used to represent the sequence, $\Re$, and mapping between the

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

variation, position, sequence and replet for the $\Phi\alpha$ replets. See variation table 4 and indirection table 5.

Further, "FIGS. 9A to 9C present "snapshots" of the variables used in the pseudo-code algorithm presented in FIG. 7 at the various stages in the algorithm when the sequence (seq3) is reconstructed from the data-structure. FIG. 9A is obtained as result of the execution of Step 820 of the algorithm, as described above. FIGS. 9B and 9C depict the values that each variable in Step 830 takes and the iteration at which those values were obtained. FIG. 9C represents Step 840 of the algorithm, in which the complete rebuilt sequence (seq3) is output as result." Present application, paragraph 99.

B. Amendments to claims herein to overcome the rejection under 35 USC 103(a)

Claim 14 is herein amended to more particularly point out aspects of the invention as described herein above. The amendment to the claims present no new matter, since the original application provides support, as explained herein above. Note Table 3 and FIG. 6, in particular.

Specifically, claim 14 now recites "for each replet, a match-set data structure comprising a sequence identification to identify a subsequence where a match occurs and offset information to determine a position within the sequence where the matching subsequence of the sequence is located, wherein the offset information comprises a first and second position parameter, the first position parameter denoting a location in the sequence and the second position parameter denoting an offset from the location" and "wherein one of the subsequences is matched by a certain plurality of the replets and the method further includes:

viii) the computer system generating or receiving a selection of one of the certain plurality of replets, wherein any non-selected one of the certain plurality of replets is deemed a redundant replet and the representing in step vii) is responsive to the selected one of the certain plurality of replets but not responsive to any redundant replet;

ix) the computer system generating or receiving a selection of one of the replets deemed a redundant replet in step viii); and

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

x) the computer system representing and presenting the sequence, wherein the representing in step x) is responsive to the one of the replets indicated by the selection in step ix) instead of the one of the replets indicated by the selection in step viii) and wherein the computer system performs the representing in step x) by reference to the first and second parameters for the one of the replets indicated by the selection in step ix), such that the computer system performs the representing in step x) without repeating the generating of position parameters performed in step iii). " The references do not teach or suggest this, neither alone, nor in combination.

C. Analysis of the cited references with regard to the present claims

A previous Office action, dated July 9, 2008, compared Taylor to what was then claimed regarding storing two position-match entries for each replet. Specifically that Office action cited Taylor, Fig. 4 and Sec. 4.1.2 regarding storing data representing amino acid alignment for at least two positions, and argued on page 8 that this shows storing two position-match parameters for recording position data of the replet. However, storing data representing amino acid alignment for two positions, does not teach or suggest "for each replet, a match-set data structure comprising a sequence identification to identify a subsequence where a match occurs and offset information to determine a position within the sequence where the matching subsequence of the sequence is located, wherein the offset information comprises a first and second position parameter, the first position parameter denoting a location in the sequence and the second position parameter denoting an offset from the location," as claimed.

Schwartz teaches a tool for comparing two DNA sequences to determine similarity of genes (such as between two related species). Schwartz, p. 577, abstract. The user submits two DNA sequence files to the Schwarz tool and is presented with several output plots that identify percent identity and aligning segments (match-sets). Schwartz, p. 578, col. 1, par.'s 2-3. This does not teach or suggest disclosing multiple ways to represent a sequence depending upon which replet is chosen for a representation.
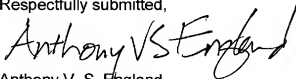
The Schwartz tool requires pre-processing of the DNA sequence data to mask repeat sequences which are "uninformative and time-consuming." Schwartz, p. 578,

Docket JP920030152US1

Appl. No.: 10/699,024
Filing Date: October 31, 2003

col. 1, par. 3. Schwartz suggests using RepeatMasker to generate "an additional file containing the coordinates of interspersed repeats in the first sequence." Schwartz, p. 578, col. 1, par. 3. The Schwartz tool does not teach or suggest allowing the user to choose among redundant replets for user control of output. Regarding significance of redundant replets, Schwartz merely teaches that "the inclusion of information about the positions and identity of repeats in the pips may aid in the discovery of informative correlations." Schwartz, p. 580, col. 2, par. 3.

REQUESTED ACTION

For the reasons explained herein above, Applicant submits that the claim 14, as amended herein, is patentably distinct. Applicant submits, further, that claims 3-6 and 15-22 are patentably distinct at least because they depend upon claim 14. Applicant, therefore, requests that all the claims be promptly allowed and the application passed to issuance.

Respectfully submitted,

Anthony V. S. England
Attorney for IBM Corporation
Registration No. 35,129
512-477-7165
a@aengland.com